

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/007,047

DATE: 12/20/2001

TIME: 14:14:26

Input Set : A:\UM-06692.ST25.txt

Output Set: N:\CRF3\12202001\J007047.raw

3 <110> APPLICANT: Mizukami, Ikuko
 4 Ross, Theodora
 5 Rao, Dinesh
 7 <120> TITLE OF INVENTION: HIP1 Cancer Markers
 9 <130> FILE REFERENCE: UM-06692
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/007,047
 C--> 11 <141> CURRENT FILING DATE: 2001-12-06
 11 <160> NUMBER OF SEQ ID NOS: 6
 13 <170> SOFTWARE: PatentIn version 3.0
 15 <210> SEQ ID NO: 1
 16 <211> LENGTH: 4534
 17 <212> TYPE: DNA
 18 <213> ORGANISM: Homo sapiens
 20 <220> FEATURE:
 21 <221> NAME/KEY: misc_feature
 22 <222> LOCATION: (4521)..(4521)
 23 <223> OTHER INFORMATION: The nucleotide "n" can be either a,t,c, or g
 26 <220> FEATURE:
 27 <221> NAME/KEY: misc_feature
 28 <222> LOCATION: (4522)..(4522)
 29 <223> OTHER INFORMATION: The nucleotide "n" can be either a,t,c, or g
 32 <400> SEQUENCE: 1
 33 ccaagcttg g taccocccgg gcagccgagg gccctgact cggctcctcg cggcgacatg 60
 35 gatcggttg ccagctccat gaagcaggtg cccaaccac tgcccaagg gctgagccgg 120
 37 cgcggggtcg gcgctgggt ggaggcggcg gagcgcgaga gcttcgagcg gactcagact 180
 39 gtcagcatca ataaggccat taatacgag gaagtggctg taaaggaaaa acacgccaga 240
 41 acgtgcatac tgggcaccca ccatgagaaa ggggcacaga ccttctggtc tgttgtaaac 300
 43 cgcctgcctc tgtctagcaa cgcagtgtc tgcctggaagt tctgccatgt gttccacaaa 360
 45 ctctctccgag atggacaccc gaacgtcctg aaggactctc tgagatacag aaatgaattg 420
 47 agtgacatga gcaggatgtg gggctacctg agcgaggggt atggccagct gtgcagcatc 480
 49 tacctgaaac tgctaagaac caagatggag taccacacca aaaatcccag gttcccaggc 540
 51 aacctgcaga tgagtgaccg ccagctggac gaggtgagag aaagtgacgt gaacaacttt 600
 53 tcccagttaa cagtggagat gtttgactac ctggagtgtg aactcaacct cttccaaaca 660
 55 gtattcaact ccttgacat gtcccgctct gtgtccgtga cggcagcagg gcagtgcgcg 720
 57 ctgcgcccg tgatccaggt catcttgagc tgcagccacc tttatgacta cactgtcaag 780
 59 cttctcttca aactccactc ctgcctccca gctgacaccc tgcaaggcca cggggaccgc 840
 61 ttcattgagc agttttacaa gttgaaagat ctgttctacc gctccagcaa cctgcagtac 900
 63 ttcaagcggc tcattcagat cccccagctg cctgagaacc caccacaact cctgcgagcc 960
 65 tcagccctgt cagaacatat cagccctgtg gtggtgatcc ctgcagaggg ctcaccccc 1020
 67 gacagcgagc cagtcctaga gaagatgac ctcattggaca tggatgcctc tcagcagaat 1080
 69 ttatttgaca acaagtttga tgacatttt ggagttcat tcagcagtga tcccttcaat 1140
 71 ttcaacagtc aaaatggtgt gaacaaggat gagaaggacc acttaattga gcgactatac 1200
 73 agagagatca gtggattgaa ggcacagcta gaaaacatga agactgagag ccagcgggtt 1260
 75 gtgctgcagc tgaagggccg cgtcagcgag ctggaagcag atctggccga gcagcagcac 1320
 77 ctgcggcagc aggcggccga cgaactgtga ttctgcggg cagaactgga cgagctcagg 1380
 79 aggcagcggg aggacaccga gaaggctcag cggagcctgt ctgagataga aaggaaagct 1440
 81 caagccaatg aacagcgata tagcaagcta aaggagaagt acagcgagct ggttcagAAC 1500

ENTERED

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/10/007,047

TIME: 14:14:26

Input Set : A:\UM-06692.ST25.txt

Output Set: N:\CRF3\12202001\J007047.raw

83	cacgctgacc	tgctgcggaa	gaatgcagag	gtgaccaaac	aggtgtccat	ggccagacaa	1560
85	gcccaggtag	atttggaacg	agagaaaaaa	gagctggagg	attcgttgga	gcgcatacagt	1620
87	gaccagggcc	agcggaagac	tcaagaacag	ctggaagtcc	tagagagctt	gaagcaggaa	1680
89	cttgccacaa	gccaacggga	gcttcagggt	ctgcaaggca	gcctggaaac	ttctgcccag	1740
91	tcagaagcaa	actgggcagc	cgagttcgcc	gagctagaga	aggagcggga	cagcctggtg	1800
93	agtggcgag	ctcataggga	ggaggaatta	tctgctcttc	ggaaagaact	gcaggacact	1860
95	cagctcaaac	tggccagcac	agaggaatct	atgtgccagc	ttgccaaaga	ccaacgaaaa	1920
97	atgcttctgg	tgggttccag	gaaggctgcg	gagcaggtga	tacaagacgc	cctgaaccag	1980
99	cttgaagaac	ctcctctcat	cagctgcgct	gggtctgcag	atcacctcct	ctccacggtc	2040
101	acatccattt	ccagctgcat	cgagcaactg	gagaaaagct	ggagccagta	tctggcctgc	2100
103	ccagaagaca	tcagtggact	tctccattcc	ataaccctgc	tggcccactt	gaccagcgac	2160
105	gccattgctc	atggtgccac	caactgcttc	agagccccac	ctgagcctgc	cgactcactg	2220
107	accgaggcct	gtaagcagta	tggcagggaa	accctcgctt	acctggcctc	cctggaggaa	2280
109	gaggaagcc	ttgagaatgc	cgacagcaca	gccatgagga	actgcctgag	caagatcaag	2340
111	gccatcggcg	aggagctcct	gcccagggga	ctggacatca	agcaggagga	gctgggggac	2400
113	ctggtggaca	aggagatggc	ggccacttca	gctgctattg	aaactgccac	ggccagaata	2460
115	gaggagatgc	tcagcaaadc	ccgagcagga	gacacaggag	tcaaatttga	ggtgaatgaa	2520
117	aggatccttg	gttgcgtgac	cagcctcatg	caagctattc	aggtgctcat	cgtggcctct	2580
119	aaggacctcc	agagagagat	tgtggagagc	ggcaggggta	cagcatcccc	taaagagttt	2640
121	tatgccaaga	actctcgatg	gacagaagga	cttatctcag	cctccaaggc	tgtgggctgg	2700
123	ggagccactg	tcatggtgga	tgacagctgat	ctggtgtgtac	aaggcagagg	gaaatttgag	2760
125	gagctaattg	tgtgttctca	tgaatttgc	gtagcacag	cccagcttgt	ggctgcatcc	2820
127	aaggtgaaag	ctgataagga	cagccccaac	ctagcccagc	tgacgcaggc	ctctcgggga	2880
129	gtgaaccagg	ccactgcgag	cggtgtggcc	tcaaccattt	ccggcaaatc	acagatcgaa	2940
131	gagacagaca	acatggactt	ctcaagcatg	acgctgacac	agatcaaacg	ccaagagatg	3000
133	gattctcagg	ttagggtgct	agagctagaa	aatgaattgc	agaaggagcg	tcaaaaactg	3060
135	ggagagcttc	ggaaaaagca	ctacgagctt	gctggtgttg	ctgagggtcg	ggaagaagga	3120
137	acagaggcat	ctccacctac	actgcaagaa	gtggttaaccg	aaaaagaata	gagccaaacc	3180
139	aacaccccat	atgtcagtg	aaatccttgt	tacctatctc	gtgtgtgtta	tttccccagc	3240
141	cacaggccaa	atccttgag	tcccaggggc	agccacacca	ctgccattac	ccagtgcgga	3300
143	ggacatgcat	gacacttcca	aagactccct	ccatagcgac	accctttctg	tttggaccca	3360
145	tggatttcca	ctgcttctta	tgggtggttg	ttgggttttt	tgggtttgtt	tttttttttt	3420
147	aagtttcaact	cacatagcca	actctcccaa	agggcacacc	cctggggctg	agtctccagg	3480
149	gccccccaac	tgtggtagct	ccagcgatgg	tgtgcccag	gcctctcggt	gctccatctc	3540
151	cgctccaca	ctgaccaagt	gctggcccac	ccagtccatg	ctccagggtc	aggcggagct	3600
153	gctgagtgc	agctttcttc	aaaaagcaga	aggagagtga	gtgcctttcc	ctcctaaagc	3660
155	tgaatcccg	cgaaaagcct	ctgtccgctt	ttacaaggga	gaagacaaca	gaaagaggga	3720
157	caagagggtt	cacacagccc	agttcccggt	acgaggctca	aaaacttgat	cacatgcttg	3780
159	aatggagctg	gtgagatcaa	caacactact	tccctgccgg	aatgaactgt	ccgtgaatgg	3840
161	tctctgtcaa	gcgggcccgc	tcccttgccc	cagagacgga	gtgtgggagt	gattcccaac	3900
163	tcctttctgc	agacgtctgc	cttgccatcc	tcttgaatag	gaagatcggt	ccaccttcta	3960
165	cgcaattgac	aaacccggaa	gatcagatgc	aattgctccc	atcagggaaq	aacctatac	4020
167	ttggtttgct	acccttagta	tttattacta	acctccctta	agcagcaaca	gcctacaaag	4080
169	agatgcttgg	agcaatcaga	acttcagggt	tgactctagc	aaagctcatc	ttctgcccgc	4140
171	gctacatcag	ccttcaagaa	tcagaagaaa	ggccaagggt	ctggactgtt	actgacttgg	4200
173	atcccaaagc	aaggagatca	tttggagctc	ttgggtcaga	gaaaatgaga	aaggacagag	4260
175	ccagcggctc	caactccttt	caqccacatg	ccccagqctc	tcqctqccct	qtggacagga	4320
177	tgaggacaga	gggcacatga	acagcttgcc	agggatgggc	agcccaacag	cacttttctt	4380
179	cttctagatg	gacccagaca	tttaagtgc	cttctgatct	tgggaaaaca	gcgtcttctt	4440

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/10/007,047

TIME: 14:14:26

Input Set : A:\UM-06692.ST25.txt

Output Set: N:\CRF3\12202001\J007047.raw

181 tctttatcta tagcaactca ttggtggttag ccatcaagca cttaggaatt cctgcagccc 4500
 183 **gggcgccgc tcgagcatgc nntagagggc ccta** 4534
 186 <210> SEQ ID NO: 2
 187 <211> LENGTH: 1475
 188 <212> TYPE: PRT
 189 <213> ORGANISM: Homo sapiens
 191 <220> FEATURE:
 192 <221> NAME/KEY: misc_feature
 193 <222> LOCATION: (1472)..(1472)
 194 <223> OTHER INFORMATION: The amino acid "X" can be any amino acid
 197 <400> SEQUENCE: 2
 199 Met Asp Arg Met Ala Ser Ser Met Lys Gln Val Pro Asn Pro Leu Pro
 200 1 5 10 15
 202 Lys Val Leu Ser Arg Arg Gly Val Gly Ala Gly Leu Glu Ala Ala Glu
 203 20 25 30
 205 Arg Glu Ser Phe Glu Arg Thr Gln Thr Val Ser Ile Asn Lys Ala Ile
 206 35 40 45
 208 Asn Thr Gln Glu Val Ala Val Lys Glu Lys His Ala Arg Thr Cys Ile
 209 50 55 60
 211 Leu Gly Thr His His Glu Lys Gly Ala Gln Thr Phe Trp Ser Val Val
 212 65 70 75 80
 214 Asn Arg Leu Pro Leu Ser Ser Asn Ala Val Leu Cys Trp Lys Phe Cys
 215 85 90 95
 217 His Val Phe His Lys Leu Leu Arg Asp Gly His Pro Asn Val Leu Lys
 218 100 105 110
 220 Asp Ser Leu Arg Tyr Arg Asn Glu Leu Ser Asp Met Ser Arg Met Trp
 221 115 120 125
 223 Gly Tyr Leu Ser Glu Gly Tyr Gly Gln Leu Cys Ser Ile Tyr Leu Lys
 224 130 135 140
 226 Leu Leu Arg Thr Lys Met Glu Tyr His Thr Lys Asn Pro Arg Phe Pro
 227 145 150 155 160
 229 Gly Asn Leu Gln Met Ser Asp Arg Gln Leu Asp Glu Ala Gly Glu Ser
 230 165 170 175
 232 Asp Val Asn Asn Phe Ser Gln Leu Thr Val Glu Met Phe Asp Tyr Leu
 233 180 185 190
 235 Glu Cys Glu Leu Asn Leu Phe Gln Thr Val Phe Asn Ser Leu Asp Met
 236 195 200 205
 238 Ser Arg Ser Val Ser Val Thr Ala Ala Gly Gln Cys Arg Leu Ala Pro
 239 210 215 220
 241 Leu Ile Gln Val Ile Leu Asp Cys Ser His Leu Tyr Asp Tyr Thr Val
 242 225 230 235 240
 244 Lys Leu Leu Phe Lys Leu His Ser Cys Leu Pro Ala Asp Thr Leu Gln
 245 245 250 255
 247 Gly His Arg Asp Arg Phe Met Glu Gln Phe Thr Lys Leu Lys Asp Leu
 248 260 265 270
 250 Phe Tyr Arg Ser Ser Asn Leu Gln Tyr Phe Lys Arg Leu Ile Gln Ile
 251 275 280 285
 253 Pro Gln Leu Pro Glu Asn Pro Pro Asn Phe Leu Arg Ala Ser Ala Leu
 254 290 295 300

RAW SEQUENCE LISTING

DATE: 12/20/2001

PATENT APPLICATION: US/10/007,047

TIME: 14:14:26

Input Set : A:\UM-06692.ST25.txt

Output Set: N:\CRF3\12202001\J007047.raw

```

256 Ser Glu His Ile Ser Pro Val Val Val Ile Pro Ala Glu Ala Ser Ser
257 305                      310                      315                      320
259 Pro Asp Ser Glu Pro Val Leu Glu Lys Asp Asp Leu Met Asp Met Asp
260                      325                      330                      335
262 Ala Ser Gln Gln Asn Leu Phe Asp Asn Lys Phe Asp Asp Ile Phe Gly
263                      340                      345                      350
265 Ser Ser Phe Ser Ser Asp Pro Phe Asn Phe Asn Ser Gln Asn Gly Val
266                      355                      360                      365
268 Asn Lys Asp Glu Lys Asp His Leu Ile Glu Arg Leu Tyr Arg Glu Ile
269                      370                      375                      380
271 Ser Gly Leu Lys Ala Gln Leu Glu Asn Met Lys Thr Glu Ser Gln Arg
272 385                      390                      395                      400
274 Val Val Leu Gln Leu Lys Gly His Val Ser Glu Leu Glu Ala Asp Leu
275                      405                      410                      415
277 Ala Glu Gln Gln His Leu Arg Gln Gln Ala Ala Asp Asp Cys Glu Phe
278                      420                      425                      430
280 Leu Arg Ala Glu Leu Asp Glu Leu Arg Arg Gln Arg Glu Asp Thr Glu
281                      435                      440                      445
283 Lys Ala Gln Arg Ser Leu Ser Glu Ile Glu Arg Lys Ala Gln Ala Asn
284                      450                      455                      460
286 Glu Gln Arg Tyr Ser Lys Leu Lys Glu Lys Tyr Ser Glu Leu Val Gln
287 465                      470                      475                      480
289 Asn His Ala Asp Leu Leu Arg Lys Asn Ala Glu Val Thr Lys Gln Val
290                      485                      490                      495
292 Ser Met Ala Arg Gln Ala Gln Val Asp Leu Glu Arg Glu Lys Lys Glu
293                      500                      505                      510
295 Leu Glu Asp Ser Leu Glu Arg Ile Ser Asp Gln Gly Gln Arg Lys Thr
296                      515                      520                      525
298 Gln Glu Gln Leu Glu Val Leu Glu Ser Leu Lys Gln Glu Leu Ala Thr
299                      530                      535                      540
301 Ser Gln Arg Glu Leu Gln Val Leu Gln Gly Ser Leu Glu Thr Ser Ala
302 545                      550                      555                      560
304 Gln Ser Glu Ala Asn Trp Ala Ala Glu Phe Ala Glu Leu Glu Lys Glu
305                      565                      570                      575
307 Arg Asp Ser Leu Val Ser Gly Ala Ala His Arg Glu Glu Glu Leu Ser
308                      580                      585                      590
310 Ala Leu Arg Lys Glu Leu Gln Asp Thr Gln Leu Lys Leu Ala Ser Thr
311                      595                      600                      605
313 Glu Glu Ser Met Cys Gln Leu Ala Lys Asp Gln Arg Lys Met Leu Leu
314                      610                      615                      620
316 Val Gly Ser Arg Lys Ala Ala Glu Gln Val Ile Gln Asp Ala Leu Asn
317 625                      630                      635                      640
319 Gln Leu Glu Glu Pro Pro Leu Ile Ser Cys Ala Gly Ser Ala Asp His
320                      645                      650                      655
322 Leu Leu Ser Thr Val Thr Ser Ile Ser Ser Cys Ile Glu Gln Leu Glu
323                      660                      665                      670
325 Lys Ser Trp Ser Gln Tyr Leu Ala Cys Pro Glu Asp Ile Ser Gly Leu
326                      675                      680                      685
328 Leu His Ser Ile Thr Leu Leu Ala His Leu Thr Ser Asp Ala Ile Ala

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/007,047

DATE: 12/20/2001

TIME: 14:14:26

Input Set : A:\UM-06692.ST25.txt

Output Set: N:\CRF3\12202001\J007047.raw

```

329      690      695      700
331 His Gly Ala Thr Thr Cys Leu Arg Ala Pro Pro Glu Pro Ala Asp Ser
332 705      710      715      720
334 Leu Thr Glu Ala Cys Lys Gln Tyr Gly Arg Glu Thr Leu Ala Tyr Leu
335      725      730      735
337 Ala Ser Leu Glu Glu Gly Ser Leu Glu Asn Ala Asp Ser Thr Ala
338      740      745      750
340 Met Arg Asn Cys Leu Ser Lys Ile Lys Ala Ile Gly Glu Leu Leu
341      755      760      765
343 Pro Arg Gly Leu Asp Ile Lys Gln Glu Glu Leu Gly Asp Leu Val Asp
344      770      775      780
346 Lys Glu Met Ala Ala Thr Ser Ala Ala Ile Glu Thr Ala Thr Ala Arg
347 785      790      795      800
349 Ile Glu Glu Met Leu Ser Lys Ser Arg Ala Gly Asp Thr Gly Val Lys
350      805      810      815
352 Leu Glu Val Asn Glu Arg Ile Leu Gly Cys Cys Thr Ser Leu Met Gln
353      820      825      830
355 Ala Ile Gln Val Leu Ile Val Ala Ser Lys Asp Leu Gln Arg Glu Ile
356      835      840      845
358 Val Glu Ser Gly Arg Gly Thr Ala Ser Pro Lys Glu Phe Tyr Ala Lys
359      850      855      860
361 Asn Ser Arg Trp Thr Glu Gly Leu Ile Ser Ala Ser Lys Ala Val Gly
362 865      870      875      880
364 Trp Gly Ala Thr Val Met Val Asp Ala Ala Asp Leu Val Val Gln Gly
365      885      890      895
367 Arg Gly Lys Phe Glu Glu Leu Met Val Cys Ser His Glu Ile Ala Ala
368      900      905      910
370 Ser Thr Ala Gln Leu Val Ala Ala Ser Lys Val Lys Ala Asp Lys Asp
371      915      920      925
373 Ser Pro Asn Leu Ala Gln Leu Gln Gln Ala Ser Arg Gly Val Asn Gln
374      930      935      940
376 Ala Thr Ala Gly Val Val Ala Ser Thr Ile Ser Gly Lys Ser Gln Ile
377 945      950      955      960
379 Glu Glu Thr Asp Asn Met Asp Phe Ser Ser Met Thr Leu Thr Gln Ile
380      965      970      975
382 Lys Arg Gln Glu Met Asp Ser Gln Val Arg Val Leu Glu Leu Glu Asn
383      980      985      990
385 Glu Leu Gln Lys Glu Arg Gln Lys Leu Gly Glu Leu Arg Lys Lys His
386      995      1000      1005
388 Tyr Glu Leu Ala Gly Val Ala Glu Gly Trp Glu Glu Gly Thr Glu
389      1010      1015      1020
391 Ala Ser Pro Pro Thr Leu Gln Glu Val Val Thr Glu Lys Glu Ser
392      1025      1030      1035
394 Gln Thr Asn Thr Pro Tyr Val Ser Val Asn Pro Cys Tyr Leu Ser
395      1040      1045      1050
397 Arg Val Cys Tyr Phe Pro Ser His Arg Pro Asn Pro Trp Ser Pro
398      1055      1060      1065
400 Arg Gly Ser His Thr Thr Ala Ile Thr Gln Cys Arg Gly His Ala
401      1070      1075      1080

```

VERIFICATION SUMMARY

DATE: 12/20/2001

PATENT APPLICATION: US/10/007,047

TIME: 14:14:27

Input Set : A:\UM-06692.ST25.txt

Output Set: N:\CRF3\12202001\J007047.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:183 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:478 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2